7057-19

U.S. ARMY INTELLIGENCE CENTER AND SCHOOL SOFTWARE ANALYSIS AND MANAGEMENT SYSTEM

ALGORITHM ANALYSIS

TECHNICAL MEMORANDUM NO. 10

MARC

Mathematical Analysis Research Corporation

11 April 1985

National Aeronautics and
Space Administration

[illegible] LABORATORY/
[illegible] of Technology
[illegible], California

JPL D-[illegible]
ALGO_[illegible]_0024

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ALGO-PUB-0024	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Technical Memo 10, 'Algorithm Analysis'		5. TYPE OF REPORT & PERIOD COVERED FINAL
		6. PERFORMING ORG. REPORT NUMBER D-4298
7. AUTHOR(s) Mathematical Analysis Research Corp. (MARC)		8. CONTRACT OR GRANT NUMBER(s) NAS7-918
9. PERFORMING ORGANIZATION NAME AND ADDRESS Jet Propulsion Laboratory, ATTN: 171-209 California Institute of Technology 4800 Oak Grove, Pasadena, CA 91109		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS RE 182 AMEND #187
11. CONTROLLING OFFICE NAME AND ADDRESS Comamder, USAICS ATTN: ATSI-CD-SF Ft. Huachuca, AZ 85613-7000		12. REPORT DATE 11 Apr 86
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Commander, USAICS ATTN: ATSI-CD-SF Ft. Huachuca, AZ 85613-7000		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE NONE
16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Dissemination		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Prepared by Jet Propulsion Laboratory for the US Army Intelligence Center and School's Combat Developer's Support Facility.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Fix Estimation; Error Budget, Analysis of Error Interaction; Simulation; Error Ellipse; Bias, Type I and II Errors		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report lists known problems associated with fix estimation and emitter identification. In many cases a list of known errors associated with the algorithm itself is appended. Finally, MARC described its simulation model and its analytic work in ~~this~~ area. PAP is described. Keywords:		

7057-19

U.S. ARMY INTELLIGENCE CENTER AND SCHOOL
Software Analysis and Management System

Algorithm Analysis

Technical Memorandum No. 10

11 April 1986

Author:

Michael W. Rennie

MARC
Mathematical Analysis Research Corporation

Approval:

James W. Gillis, Subgroup Leader
Algorithm Analysis Subgroup

Edward J. Records, Supervisor
USAMS Task

Concur:

A. F. Ellman, Manager
Ground Data Systems Section

Fred Vote, Manager
Advanced Tactical Systems

JET PROPULSION LABORATORY
California Institute of Technology
Pasadena, California

JPL D-4296

PREFACE

The work described in this publication was performed by the Mathematical Analysis Research Corporation (MARC) under contract to the Jet Propulsion Laboratory, an operating division of the California Institute of Technology. This activity is sponsored by the Jet Propulsion Laboratory under contract NAS7-918, RE182, A187 with the National Aeronautics and Space Administration, for the United States Army Intelligence Center and School.

EXECUTIVE SUMMARY

This Technical Memorandum was prepared to satisfy questions about what Algorithm Analysis is all about and what are some of the past results and fruitful areas for continued study. This report was requested by CDSF personnel during FY-86 and used FY-86 funds.

This report, although late in the sequence, should be read as early as possible and by the general reader who may not be interested in all the knitty-gritty mathematical/engineering details.

Algorithm Analysis

Introduction

An algorithm should be judged against other algorithms in terms of its ability to meet objectives using available resources. For the algorithms that have been studied, the objectives are understood to be:

(1) Identification of signals, emitters and/or units.
(2) Location of the location estimate. (Accuracy is indicated by means of error ellipses.)

An overview starts with the objectives given above. A more detailed examination is given in the following section of the memo:

I. TOP LEVEL ISSUES - This is a top level definition of issues that analysis should address.

II. THE 'COMMON STRUCTURE' OF THE ALGORITHMS

III. POTENTIAL SOURCES OF PROBLEMS - (Organized according to the structure defined in section II.) This is needed to define the issues for which algorithms are examined. Needless to say the list given here is a first cut.

IV. ISSUES AND ALGORITHMS EXAMINED TO DATE

V. AN INTEGRATED MODEL - a simulation model with replaceable 'modules' is being developed. The potential and the limitations of this or other models is discussed.

VI. AREAS OF INTEREST

Before proceeding to the topics described above there is one observation which affects all sections of this report. The objective of identification much different from the objective of locating an emitter. This is primarily owing to the fact that identification is primarily discrete (finite possibilities) whereas location is continuous (continous possibilities). This makes identification errors more catastrophic. There is not even any measure of the confidence in an identification reported or used (that has been seen.) Identification issues are less easily modeled and hence are an obvious potential concern.

I. TOP LEVEL ISSUES

An outline of top level issues is given following at the end of this section. In the paragraphs that follow the relation of outline to the two objectives given in the Introduction is explained.

The first of the two objectives given in the Introduction is identification. Identification issues occur at all levels of the algorithm and there may be applications that would use these intermediate results. Thus a number of identification issues are outlined below under the heading misclassification. In some cases, however, a particular type of misidentification may be of more interest in terms of the impact it has as a source of error in other algorithms.

The second of the two objectives given in the Introduction is location. Location issues are easier to characterize and hence a number of sub-headings apply. These are:

Location Bias - does the algorithm locate correctly on average?
Speed of Convergence - number and accuracy of LOB's needed to achieve a given accuracy on average.
Accuracy of Elipse Size - so that you know when you have converged sufficiently.

Two other incidental headings are given in the outline but without much elaboration.

A. Misclassification (erroneous identification decisions)
   (1) Types of Misclassification
      a) Identifying an emitter that does not exist or failing to identify one that does exist.
      b) Failing to associate an LOB with the proper emitter and associating the ellipse with an erroneous emitter (during the fix process.)
      c) Failing to associate an ellipse with the proper emitter and associating the ellipse with an erroneous emitter (during the ellipse combination process.)
      d) Misclassificaiton of emitter or unit type.
      e) Data transmission, editing and censoring related errors.
   (2) Classification Confidence
      a) How much is there?
      b) Can it be used by algorithms?

B. Location Bias (does the algorithm locate correctly on average)
   (1) Types of Bias
      a) Bias induced by the fix process
      b) Bias induced by the ellipse combination process
   (2) Interaction of Bias with other problems
      a) The impact of bias on classification
      b) The impact of classification on bias (if any)
   (3) Bias as a function of sample size (and other sensor baseline data)

C. Speed of Convergence (number and accuracy of LOB's needed to achieve a given accuracy on average)
   (1) Comparative rate at which error ellipses shrink.
   (2) 'True' rate at which the estimate converges.

(3) Stopping criteria.

D. Accuracy of Ellipse Size (so that you know when you have converged)
   (1) Almost every consideration affects ellipse size to some extent. Major sources of variation should be modeled if possible.
   (2) Interaction of Ellipse Size and Misclassification

E. Anomalous Cases (Worst case analysis)

F. Computational Complexity (ability to process the data)

II. <u>THE 'COMPUTATIONAL STRUCTURE' OF THE ALGORITHMS</u>

A. SIGINT Sensors
   (1) Bearings
      a) Determination of the direction of a bearing
      b) Identification of signal characteristics
   (2) Emitter Refinement
      a) Identification of the set of bearings belonging to a particular emitter.
      b) Determination of a position estimate
      c) Determination of an error ellipse
      d) Characterization of signal parametrics

B. JINTACCS
   (1) Selected data is transmitted
   (2) Precision is limited

C. SIM's (Sensor Interface Modules)
   (1) Normalize EEP's is possible
   (2) Edit Data

D. Pre-Sorting
   It is unclear what this does. It may create conditional probabilities that are not adequately modeled.

E. Test for Ellipse Combination (Normally Chi-square)
   (1) Test to determine if data should be identified with an existing emitter or as a new emitter.
   (2) Test for movement.

F. Ellipse Combination

G. Node Identification

III. <u>POTENTIAL SOURCES OF PROBLEMS</u>

1. Bearing Direction

   a) Calibration
   b) Bearing 'Quality' Problems

2. Bearing Signal Characteristics

(MARC has not worked in this area and hence has no list here)

3. Initial Establishment of Emitter Existence

   a) Nearby emitters
   b) Wild bearing contaminated environment
   c) Slow to identify a new emitter
   d) Too quick in identifying a new emitter

4. Identification of the Set of Bearing Belonging to a particular Existing Emitter

   a) Nearby emitters
   b) Wild bearing contaminated environment
   c) Bearing signal characteristic issues

5. The position Estimate From the Fix Algorithm

   a) Bias (and bias as a function of sample size)
   b) Incoming EEP's have correlated error (such as True North error)
   c) Speed of convergence

6. Error Ellipse From the Fix Algorithm is the Wrong Size

   a) Error ellipse size is small compared to bias
   b) Sensor location error is not accounted for
   c) Wild bearings are not accounted for
   d) The method of selecting the EEP truncates the incoming distribution
   e) Incoming EEP's are not normally ditributed
   f) Calibration
   g) Algorithm does not cause emitter to land in the EEP 95% of the time (even in benign simulations)

7. Characterization of Signal Parametrics
   (MARC has not worked in this area and hence has no list here)

8. JINTACs
   (MARC has not worked in this area and hence is speculating)

   a) Is data garbled or inappropriately handled with a frequency to make this issue a concern?
   b) Is the truncation significant?

9. SIMs
   a) Inappropriate Normalization

10. Pre-Sorting
    (MARC has not worked in this area and hence has no list)

11. Test for Ellipse Combination

a) Is this an error trade-off situation rather than a simple test enviroment?
b) Is the appropriate if the ellipse is based in the F-test?
c) Impact of bias on the test
d) Impact of incorrect ellipse size.

12. Ellipse Combination

a) Impact of incoming errors such as bias
b) Bias contribution of the combination algorithm (particularly if large numbers of 'large' ellipses are used
c) Repeated or duplicated data

13. Node Identification
(MARC has not worked in this area and hence has no list).

IV. ISSUES AND ALGORITHMS EXAMINED TO DATE

This section will be organized around the issues fo section III. The work done discussed in this section only included work done by MARC.

No Work Done To Date -

Bearing direciton, bearing signal characteristics, characterization of signal parametrics, JINTACs, Pre-sorting, node identification

Work Done -

*** Initial Establishment of Emitter Existence

MARC has examined the Jackknifing technique used by Guardrail briefly. This examination included a comparison of the statistical test used in this technique with the tests for adding new LOB's to refine location and with the test for ellipse combination. This comparison was an offshoot of a study of similarity of LOB combination and ellipse combination.

*** Identification of the Set of Bearing Belonging to a Particular Existing Emitter

There are two aspects to this problem. Rejection of a 'true' LOB and acceptance of a 'false' LOB. The first work was an initial look 1-dimensional acceptance tests in terms of both problems. Subsequent work has separated these two issues. Acceptance of a false LOB was studied briefly owing to an inquiry from Fort Huachuca and a wild bearing report was issued. This report was primarly concerned with the impact on ellipse size. This work led MARC to include a provision for a second emitter in its simulation model. This approach has not been pursued recently but is on MARC's list of things to do eventually. More recently interest in a comparison of ellipse combination versus LOB combination caused MARC to examine the issues from a different point of view. This examination is not complete. MARC is in the process of examining the technique used by Guardrail in terms of comparing LOB combination to ellipse combination. Several other of MARC's projects will cause MARC to examine this issue from other points of view. The motivation for this will be

i) to determine if or how the truncation of the normal distribution of errors implicit in the test for acceptance affects our study of bias.
ii) to consider the addition of an option to perform this test in our simulation model (this would be a major change as our program currently does not specify an order for LOB's.)

*** The Position Estimate From the Fix Algorithm

a) Bias (and bias as a function of sample size)

This is a major area of investigation by MARC and will be the subject of MARC's next large report.

b) Incoming EEP's have correlated error (such as True North error.)

This is 'scheduled.'

c) Speed of convergence

Not explicitly studied. Not that an initial study of what each method is claiming from its rules for error ellipse determination might not be difficult. Some observations are very simple. The F-test based weighted perpendicular method (similar to Guardrail) starts slower than the chi-square but has essentially caught up after about 20 LOBs. Actual convergence is a different issue. (In this case the two methods would be identical as they compute the same point estimates.)

*** Error Ellipse From the Fix Algorithm Is the Wrong Size

a) Error ellipse size is small compared to bias

This issue is being studied both analytically and in terms of MARC's simulation program.

b) Sensor location error is not accounted for

An adjustment scheme was worked out but it requires knowledge of the probability of a wild bearing. The first step towards incorporation into the simulation program has been taken.

d-g) The other sub-issues have been studied.

*** Inappropriate Normalization

This has been the subject of two short investigations MARC performed at the request of JPL and Fort Huachuca. One is documented in a MARC memo entitled "Results of Inappropriate EEP Normalization Methods in Correlation" and dated 17 September 1985. The other coincides with the writing of this report.

*** Test for Ellipse Combination

MARC studied general properties of the combination test in the report "Testing and Combination of Confidence Ellipses: A Geometric Analysis."

a) Is this an error trade-off situation rather than a simple test environment.

This issue was studied briefly early on when MARC was still learning the algorithm. MARC's simulation models are almost sophisticated enough now to revisit the issue.

b) Is the test appropriate if the ellipse is based on the F-test?

Not yet studied but on MARC's list.

c) Impact of bias on the test

Studied early on with one dimensional example and now being studied at a more realistic level as part of the current bias study.

d) Impact of incorrect ellipse size

This is a simple issue and does not much require study. it should probably be documented, however. MARC has only examined specific examples to date.

*** Ellipse Combination

MARC studied general properties of the combination in the report "Testing and Combination of Confidence Ellipses: A Geometric Analysis."

a) Impact of incoming errors such as bias

b) Bias contribution of the combination algorithm (particularly if large numbers of 'large' ellipses are used)

MARC's analysis has come out of its study of the similarity of ellipse combination and LOB combination.

c) Repeated or duplicated data

This makes the resultant ellipse too small. The amount too small is approxiamtely proportional to the square root of the average repetition. If the repetition is sporadic, then the effect would probably be somewhat less than this.

V. AN INTEGRATED MODEL

The issues that have been raised in this memo need not be the ones a model is built around. It is necessary that the list of issues be developed before the model is built, however. As an example, MARC's current computer model was not designed to give an order to LOB's because it wasn't needed. Adding it to check LOB selection criteria will require adding order logic to all other routines. The bigger the model the more of the code that must be redesigned. Fortunately, in our example no one else has to modify their code to accomdate our changes. Almost no one at least. JPL has been translating our simulation programs. Their job would be easier if we could have planned the program instead of adding features as we think of them.

Another aspect of generation of an integrated model such as the one desired is that the replaceable parts of the model must be constructed to that they have the same input and output. This requires an understanding of the range of possibilities for each replaceable portion of the model. The use of real breaking points helps. An example here that MARC has not used is the JINTAC message. If MARC's model is going to be used extensively it is probably important that the JINTAC message be incorporated into the model soon. (It hasn't been incorporated to date in part because MARC has not studied the JINTAC message.)

As currently designed MARC's model can be used to address a number of issues. The sooner the full scope is defined the smoother the development will be. This same hurdle must be crossed if it were to be decided to start development along the lines of this memo and exploration of the range of algorithms at each stage is logical at this point.

Simulation should not be considered as an alternative to analysis. Both simulation and analysis should support each other. It should be possible to design a model for investigation of algorithms in relation to well defined issues. This same model could also be used for illustration of algorithm behavior. Small variations might make the model applicable to intelligence assesment.

VI. AREAS OF INTEREST TO MARC

MARC wishes to:

(1) Continue working in the areas listed in section IV as areas that MARC has already done work in.

(2) Expand our simulation model.

(3) Analyze or develop identification algorithms. (Since the low end identification algorithms may more naturally fall onto the hardware side MARC's work might be restricted to fusion or cross-correlation.)

These criteria are less important to MARC than a general desire to work on problems where our talents can be put to best use.